

## INTISARI

### **Analisis Sentimen Konsumen Review pada Marketplace Menggunakan Metode Word2Vec dan BERT (Bidirectional Encoder Representations From Transformers)**

Oleh

Dimas Midyan Alam

21/486709/PPA/06252

Analisis sentimen adalah sebuah proses yang memahami, mengekstraksi, dan mengolah data teks secara otomatis untuk menemukan jenis sentimen pada teks tersebut, merupakan solusi mengatasi salah satu masalah dalam bidang usaha untuk mengelompokkan opini atau review menjadi opini positif atau negatif secara otomatis. Analisis sentimen berguna untuk memudahkan pengguna pada proses memahami sentimen sehingga dapat melakukan penentuan keputusan pada suatu objek.

Pada penelitian ini memanfaatkan metode *Word2Vec* dan *Bidirectional Encoder Representation from Transformer* (BERT) sebagai metode *Word Embedding* untuk merepresentasikan kata menjadi vektor. Metode *Word Embedding* digunakan untuk merepresentasikan kata-kata dalam bentuk matematis, merupakan sebuah algoritma untuk mempelajari posisi kedekatan semantik antar kata dari sebuah teks masukan. Hasil dari ekstraksi fitur tersebut akan diklasifikasikan menggunakan metode *Long Short Term Memory* (LSTM), untuk mengklasifikasikan opini atau review kedalam label positif dan negative. Parameter evaluasi yang digunakan dalam penelitian ini adalah akurasi, presisi, recall, dan F1 Score pada klasifikasi dengan menggunakan metode *Word2Vec*, *BERT* dan tanpa menggunakan *Word Embedding*. Pada penelitian yang dilakukan menggunakan 5017 untuk data sentimen negative dan 5005 untuk data sentiment positive, dengan total dataset berjumlah 10022 diperoleh nilai akurasi analisis sentiment tanpa menggunakan word embedding adalah 88%, sedangkan yang menggunakan word2vec 92%, dan pada metode BERT menghasilkan nilai akurasi sebesar 95%.

Kata Kunci: Sentimen Analisis, Word2Vec, BERT, Word Embedding, LSTM

## **ABSTRACT**

### **Sentiment Analysis Consumer Review on the Marketplace Using Word2Vec and BERT Methods (Bidirectional Encoder Representations From Transformers)**

Oleh

Dimas Midyan Alam

21/486709/PPA/06252

Sentiment analysis is a process that understands, extracts, and processes text data automatically to find the type of sentiment in the text. It is a solution to one of the problems in the field of enterprise to group opinions or reviews into positive or negative opinions automatically.

In this study, Word2Vec and Bidirectional Encoder Representation from Transformer (BERT) methods were used as word embedding methods to represent words into vectors. The results of the feature extraction will be classified using the Long Short Term Memory (LSTM) method to classify opinions or reviews into positive and negative labels. The evaluation parameters used in this study are accuracy, precision, recall, and F1 score on classification using Word2Vec, BERT methods, and without using Word Embedding. In a study conducted using 5017 for negative sentiment data and 5005 for positive sentiment data, with a total dataset of 10022, sentiment analysis accuracy values without using word embedding were 88%, whereas using word2vec was 92%, and the BERT method yielded accurate values of 95%.

**Keywords:** Sentiment Analysis, Word2Vec, BERT, Word Embedding, LSTM